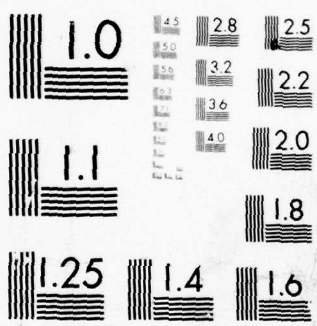
AD-A072 407 SOUTHERN METHODIST UNIV DALLAS TX DEPT OF OPERATIONS--ETC F/G 12/1
COMPUTER GENERATION OF BIVARIATE GAMMA RANDOM VECTORS.(U)
JUL 79 B W SCHMEISER, R LAL N00014-77-C-0425
UNCLASSIFIED OR-79009 NL

1 OF 1

AD
A072407

END
DATE
FILMED
9-79
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-

AD A072407

LEVEL II

12 3S.

9 Technical Rep[redacted]

6 COMPUTER GENERATION OF BIVARIATE

GAMMA RANDOM VECTORS.

12 37 p'

10 Bruce W. Schmeiser

Ram Lal

14 OR-79009

11 Jul 79

Department of Operations Research
Southern Methodist University
Dallas, Texas 75275

This research was supported by Office of Naval Research Contract

15 N00014-77-C-0425 [redacted]

411 321

-A-

ABSTRACT

The use of bivariate gamma distributions in simulation modeling is considered. A family of algorithms is given, any member of which can be used to generate random bivariate vectors having any gamma marginal distributions and any theoretically possible correlation, both positive and negative. Computational considerations are discussed, including the selection of parameters to provide regression curves consistent with the modeler's needs. A modification which is easier to implement and provides most, but not all, correlations is also given. Finally the use of these algorithms to generate first order autoregressive time series is discussed.

Many random phenomena may be modeled as gamma random variables and many random phenomena are correlated. For example, time between "events" in the analysis of nulcear power plant safety, population characteristics such as income and age, and species' lifetimes are all nonnegative random variables which may be modeled by the gamma distribution in some cases. In addition, these variables are often subject to the same exogenous variables, causing them to be either positively or negatively correlated. For example, two power plants may be affected by the same earthquake causing positive correlation, age and income may be correlated, and a higher than average rainfall may be helpful to one species and harmful to another species, causing negative correlation in their lifetime.

While it would then appear that bivariate or multivariate gamma models should appear frequently in the literature, this does not seem to be the case. Instead, the multivariate normal distribution is commonly used when correlated random variables are appropriate, since multivariate normal properties are well known and random vectors can be generated relatively easily (Scheuer and Stoller [34], Deik [9] and Schmeiser and Ali [35]).

In the unidimensional case, the gamma distribution is a generalization of the normal in that the normal distribution is a limiting distribution of the gamma as the shape parameter goes to infinity. Thus in the multidimensional case, a multivariate gamma model is more general than the multivariate normal model, thereby allowing more general modeling ability.

In this paper a bivariate gamma distribution is developed in the form of a family of algorithms capable of generating random vectors

possessing gamma marginal distributions and specified correlation coefficient ρ. In particular the marginal distributions have density functions

$$f_i(x) = ((x/\beta_i)^{\alpha_i - 1} \exp(-x/\beta_i) \,/\, (\beta_i \,\Gamma(\alpha_i))$$

$$\text{for } x > 0,\ \alpha_i > 0,\ \beta_i > 0,\ i = 1, 2$$

The correlation ρ does not depend upon the scaling parameters β_1 and β_2, so we assume $\beta_1 = \beta_2 = 1$ in the rest of this paper. The desired scaling can be achieved by multiplying the generated variates by their respective scaling parameters, as is done in the algorithm in the Appendix.

The results in this paper differ from previous work, discussed in Section 1, in that the algorithms developed here do not require integer shape parameters α_i, nor equal shape parameters, and they provide variates having any theoretically possible correlation ρ. As seen in later Section 1, most existing methods allow only positive correlations, and then not all possible positive correlations.

Thus the range of possible correlations for given shape parameters α_1 and α_2 is of interest. For the bivariate normal $\rho \in [-1, 1]$. However for gamma marginal distributions, not all correlations are consistent with particular shape parameter values. Figure 1 shows the obtainable correlations as a function of α_2, given $\alpha_1 = 1$ and 5. Correlations inconsistent with the specified shape parameters are shaded. Note that only when $\alpha_1 = \alpha_2$ is it possible for $\rho = 1$. Likewise $\rho = -1$ is not possible except in the limit as α_1 or α_2 tend to infinity. The dashed lines are discussed in Section 1.

Figure 1 About Here

The maximum and minimum possible correlations, given in Moran [31], occur when $X_2 = F^{-1}_{\alpha_2}(F_{\alpha_1}(X_1))$ and $X_2 = F^{-1}_{\alpha_2}(1-F_{\alpha_1}(X_1))$, respectively, where $F_\alpha(x)$ and $F^{-1}_\alpha(u)$ are the cumulative distribution function (cdf) and inverse cdf, respectively, of the gamma distribution with shape parameter α. These relationships are symmetric and can alternatively be expressed by interchanging all subscripts. In Figure 2, these relationships are illustrated for a specific value $X_1 = x_1$, with $X_2 = x_2$ being the corresponding value when correlation is maximized and $X_2 = x_2'$ being the corresponding value when the correlation is minimized.

Figure 2 About Here

---

1. EXISTING METHODS FOR BIVARIATE GAMMA GENERATION

There are several existing methods for generating random bivariate gamma vectors, although each has a limitation which makes it less general than the algorithms developed in Section 2.

Trivariate reduction, developed by Arnold [2], can be used to generate variates having positive correlation with $\rho < \min(\alpha_1, \alpha_2)/(\alpha_1\alpha_2)^{1/2}$. Letting "gamma (α, β)" denote the gamma distribution with parameters α and β, the trivariate reduction algorithm proceeds as follows for given parameter values α_1, α_2, and ρ.

<u>ALGORITHM GTVR</u>

1. Generate $Y_1 \sim \text{gamma}(\alpha_1 - \rho(\alpha_1\alpha_2)^{1/2}, 1)$
2. Generate $Y_2 \sim \text{gamma}(\alpha_2 - \rho(\alpha_1\alpha_2)^{1/2}, 1)$
3. Generate $Y_3 \sim \text{gamma}(\rho(\alpha_1\alpha_2)^{1/2}, 1)$
4. $X_1 \leftarrow Y_1 + Y_3$
5. $X_2 \leftarrow Y_2 + Y_3$

As noted earlier, X_1 and X_2 can be multiplied by β_1 and β_2, respectively, to obtain any desired scaling. In algorithm GTVR, Y_3 is a common component of both X_1 and X_2, thus inducing positive correlation. The dashed lines in Figure 1 indicate the boundaries of correlation obtainable with trivariate reduction. When a positive, but not extreme, correlation is needed, trivariate reduction is an excellent algorithm. Fishman [11] suggests this method for generating correlated gamma variates.

Ronning [33] generalized trivariate reduction, although no specific algorithms are given. Rather a general framework is discussed, from which it appears possible to create algorithms which would be able to obtain any positive correlation ρ. In addition, he considered the n dimensional case.

Composition, or probabability mixing, provides a very general algorithm for generating random bivariate vectors, in that all possible parameter values may be obtained, including negative correlations. The algorithm proceeds as follows:

ALGORITHM GCOMP

1. Generate $U \sim U(0,1)$
2. If $U > \rho/C$, go to step 5.
3. Generate $X_1 \sim \text{gamma}(\alpha_1, 1)$
4. Generate $X_2 \sim \text{gamma}(\alpha_2, 1)$ and go to step 9.
5. Generate $V \sim U(0,1)$ (or $V \leftarrow (U*C - \rho)/(C - \rho)$)
6. $X_1 \leftarrow F^{-1}_{\alpha_1}(V)$
7. If $\rho < 0$, $V \leftarrow 1 - V$
8. $X_2 \leftarrow F^{-1}_{\alpha_2}(V)$
9. Deliver (X_1, X_2)

Here C is the maximum or minimum possible correlation given α_1 and α_2. C is calculated using equation (3) in Section 2 with $H_1 = \alpha_1$, $H_2 = \alpha_2$ and $\gamma = 0$. The implementation of algorithm GCOMP could proceed in many ways, with the given logic being an example. The underlying idea is to generate an independent vector (X_1, X_2) with probability ρ/C and to generate a vector (X_1, X_2) having correlation C otherwise.

The most severe problem with the composition approach is the statistical properties of the generated variates. Figure 3 is a scatter plot of 1000 random generated vectors with $\alpha_1 = 10$, $\alpha_2 = 10$, and $\rho = -.5$. The marginal distributions and the correlation are as desired. However, the points fall in a pattern that is inappropriate in most applications.

The pattern is that of many independent variates superimposed over many other variates lying on the curve $X_2 = F^{-1}_{\alpha_2}(1-F_{\alpha_1}(X_1))$. Examination of Figure 3 thus shows that in addition to obtaining the correct marginal distributions and correlation, some thought must be given to the pattern of vectors obtained.

Figure 3 About Here

A third approach, advocated by Kottas and Lau [25], attacks the pattern directly, by modeling the regression line $E(X_2|x_1)$ and the variance $V(X_2|x_1)$ explicitly. Their work, done in a much more general context, does not provide for specified marginal distributions or specified correlations. Thus, to a large extent, the method chosen depends upon the data available to the modeler. When the structure of the dependency between X_1 and X_2 is understood, using this understanding is an excellent approach. In related work, Hull [17] approximates the desired correlation while considering the regression line.

Moran [31], Gunst and Webster [16], Johnson and Kotz [22] and Kibble [23] discuss and reference several other multivariate gamma distributions, but all have restrictions such as $\alpha_i = \alpha_j$, $2\alpha_i$ integer, or $\rho \geq 0$. While none were developed for use in simulation, they could be useful in some cases.

There are several papers which provide methods which could be used to approximate a bivariate gamma distribution with specified correlation $\rho > 0$. Lee [26] discusses a multivariate Weibull distribution. Takahasi [39] and Durling, Owen, and Drane [10] discuss multivariate Burr distributions with $\rho > 0$. Mihram and Stacy [36] discuss a warning-time/failure-time bivariate distribution with beta and generalized four-parameter

gamma marginal distributions. Johnson [21] considers translations of bivariate normal vectors. Johnson and Ramberg [19], developed a bivariate uniform distribution.

For $\rho < 0$, few methods have been developed. Gargano and Tenenbein [13] and Johnson and Tenenbein [20] have developed bivariate uniform distributions which can be used to generate variates having negative correlations. Using rank correlation, bivariate gamma distributions can be generated by transforming the bivariate uniform variates to gamma variates via an inverse cdf transformation.

In the related area of autocorrelated sequences, two papers suggest methods for $\rho < 0$: Polge, Holliday, and Bhagavan [32] suggest a "correlation transfer" method, which provides approximate correlation and marginal distributions; and Gaver, Lavenberg, and Price [14] propose generating correlated variates with negative correlation via a Markov chain model.

2. A FAMILY OF ALGORITHMS

Developed in this section is a family of algorithms, any member of which can produce bivariate gamma vectors having any parameters α_1, α_2 and ρ.

Proposition 1. If Z, W_1, and W_2 are independent gamma random variables with shape parameters γ, δ_1, and δ_2, respectively; and U is an independent $U(0,1)$ random variable; and either $V = U$ or $V = 1 - U$; then

$$X_1 = F_{H_1}^{-1}(U) + Z + W_1$$

and (1)

$$X_2 = F_{H_2}^{-1}(V) + Z + W_2$$

are each gamma random variables, with shape parameters $\alpha_1 = H_1 + \gamma + \delta_1$ and $\alpha_2 = H_2 + \gamma + \delta_2$, respectively.

Proof: The result follows immediately from the reproducibility property of the gamma distribution and from noting that $F^{-1}(U)$ and $F^{-1}(1-U)$ are each random variables having cdf F.

Proposition 2. $\mathrm{Corr}(X_1, X_2) = [E(F_{H_1}^{-1}(U)F_{H_2}^{-1}(V)) - H_1H_2 + \gamma]/(\alpha_1\alpha_2)^{1/2}$

Proof: The proof is straightforward algebra.

Propositions 1 and 2 say that equations (1) can be used to generate (X_1, X_2) having the desired marginal distributions and correlation.

Given α_1 and α_2, any theoretically possible correlation ρ may be obtained. Examination of some important limiting cases is helpful

to understand the robustness of the algorithms:

1. Let $\gamma = \delta_1 = \delta_2 = 0$. Then the $\mathrm{Corr}(X_1, X_2)$ is maximized if $V = U$ and minimized if $V = 1 - U$, as discussed earlier.
2. Let $H_1 = 0$ or $H_2 = 0$, and $\gamma = 0$. Then X_1 and X_2 are independent random variables.

Letting the parameters range between these two extreme cases yields all possible desired correlations.

The remaining problem is to select values of the five parameters H_1, H_2, γ, δ_1 and δ_2 to obtain the desired marginal distributions and correlation. The following conditions must be satisfied:

$$\begin{aligned} &H_1 + \gamma + \delta_1 = \alpha_1 \\ &H_2 + \gamma + \delta_2 = \alpha_2 \\ &[E(F_{H_1}^{-1}(U)F_{H_2}^{-1}(V))-(H_1H_2) + \gamma] \,/\, (\alpha_1\alpha_2)^{1/2} = \rho \\ &H_1, H_2, \gamma, \delta_1, \delta_2 \geq 0. \end{aligned} \tag{2}$$

Since we are using five variables to satisfy three equality conditions, finding a set of parameter values corresponds to finding a feasible solution, rather than an optimal solution, to a nonlinear programming problem. Here δ_1 and δ_2 are slack variables.

An efficient solution procedure for determining parameter values is important, since substantial computation is required to determine whether or not conditions (2) are satisfied for given parameter values. Most of the computation is involved in calculating

$$\mathrm{Corr}(X_1, X_2) = [E(F_{H_1}^{-1}(U)\, F_{H_2}^{-1}(V))-(H_1H_2) + \gamma] \,/(\alpha_1\alpha_2)^{1/2} \tag{3}$$

since the expected value must be calculated numerically using any one of the following three integrals:

$$\int_0^1 F_{H_2}^{-1}(u) \quad F_{H_1}^{-1}(u)\ du \tag{4}$$

$$\int_0^\infty F_{H_2}^{-1}(F_{H_1}(x))\ x^{H_1} \exp(-x)\ dx \ / \ \Gamma(H_1) \tag{5}$$

$$\int_0^1 F_{H_2}^{-1}(F_{H_1}(-\ln y))\ (-\ln y)^{H_1} dy \ / \ \Gamma(H_1) \tag{6}$$

if $\rho > 0$. If $\rho < 0$, then replace $F_{H_1}^{-1}(u)$ with $F_{H_1}^{-1}(1-u)$ in equation (4), replace $F_{H_1}(x)$ with $1-F_{H_1}(x)$ in equation (5), and replace $F_{H_1}(-\ln y)$ with $1-F_{H_1}(-\ln y)$ in equation (6). Davis and Rabinowitz [8] and other numerical integration textbooks discuss various methods for evaluting these integrals. We seemed to have best results, in terms of a subjective trade off between speed and accuracy, using a 24 point Gaussian method on integral (4). Integrals (5) and (6) have the advantage of requiring only one inverse cdf evaluation, which is advantageous since $F^{-1}(u)$ is iteratively calculated from $F(x)$ in almost all published methods. (See, e.g., Best and Roberts [5]).

One way to select the parameter values is to choose the feasible values satisfying conditions (2) and to optimize by making the curves of regression $E(X_1|x_2)$ and $E(X_2|x_1)$ behave as desired. Proposition 3 gives these curves as a function of the parameters.

Proposition 3. If X_1 and X_2 are defined as in equations (1) and $\rho > 0$,

$$E(X_2|x_1) = \int_0^1 F_{H_2}^{-1}(F_{H_1}(x_1\ u))\ u^{H_1-1}\ (1-u)^{\alpha_1-1}\ /\beta(H_1,\alpha_1)\ du + \gamma x_1/\alpha_1 + \delta_2$$

and $E(X_1|x_2)$ is obtained symmetrically by interchanging all subscripts "1" and "2". If $\rho < 0$, replace $F_{H_1}(x_1 u)$ by $1-F_{H_1}(x_1 u)$.

Proof: The results follow directly given that the distribution of $F_{H_1}^{-1}(U)$ given x is beta with parameters (H_1, α_1) over the range [0, x] and the distribution of Z given x is beta with parameters (γ, α_1) over the same range.

An important special case arises when $H_1 = H_2$ and $\rho > 0$, since then

$$E(X_1 | x_2) = x_2(H_1 + \gamma)/\alpha_2 + \delta_1$$

and

$$E(X_2 | x_1) = x_1(H_2 + \gamma)/\alpha_1 + \delta_2.$$

Note that these lines are not identical, nor are they even parallel, although they may be made parallel by setting $\gamma = (\alpha_1\alpha_2)^{1/2} - H_1$. The ability to obtain nonlinear lines of regression is an important advantage of this general family of algorithms compared to the more limited trivariate reduction algorithm, in which always $H_1 = H_2 = 0$.

The generation of the bivariate vector (X_1, X_2) can be performed directly from Equations (1), with Z, W_1, and W_2 being generated by any gamma generator. Theoretically exact gamma generators are given in Ahrens and Dieter [1], Atkinson [3], Atkinson and Pearce [4], Cheng [6], Fishman [11, 12], Jöhnk [18], Greenwood [15], Kinderman and Monahan [24], Marsaglia [27], McGrath and Irving [28], Schmeiser and Lal [36], Tadikamalla [37, 38], C. S. Wallace [40], N. D. Wallace [41], and Whittaker [42]. Speed is not a factor here since the evaluation of $F^{-1}(U)$ requires orders of magnitude more time than the fastest of these algorithms. Care need be taken only to not use an algorithm with time requirements which are unbounded for some parameter values.

Considerable execution time can be saved by implementing the algorithm less directly. Rather than performing two slow evaluations of F^{-1}, the following requires only one F^{-1} evaluation:

Generate $X_1 \sim \text{gamma}(H_1)$

$U \leftarrow F_{H_1}(X_1)$

If $\rho < 0$, $U \leftarrow 1 - U$.

Generate $Z \sim \text{gamma}(\gamma, 1)$

Generate $W_1 \sim \text{gamma}(\delta_1, 1)$

Generate $W_2 \sim \text{gamma}(\delta_2, 1)$

$X_1 \leftarrow X_1 + Z + W_1$

$X_2 \leftarrow F_{H_2}^{-1}(U) + Z + W_2$

3. SELECTION OF A PARTICULAR ALGORITHM

Several criteria can be identified for selecting a particular algorithm from the family developed in Section 2:

1. Set-up time required.
2. Marginal execution time for each vector (X_1, X_2) generated.
3. Program bulk (related to lines of code and support routines needed).

and

4. Statistical properties of the generated variates.

This fourth criterion is probably the most important since the first three criteria are very similar for all algorithms in the family. Thus we chose one algorithm for implementation based on criterion 4, as measured subjectively by viewing scatter plots for many values of α_1, α_2, and ρ. The algorithm given in detail in the appendix seemed to behave as "expected". While none of the algorithms were "bad", some tended to produce a rather sharp boundary on the generated values. The algorithm selected tended to produce values which tapered off gradually, which to the authors appeared to be subjectively better. Figure 4 is a scatter plot produced by the selected algorithm.

Figure 4 About Here

The algorithm implemented in the Appendix corresponds to $\gamma = \delta_2 = 0$ and to assuming that $\alpha_1 \geq \alpha_2$. The second assumption is without loss of generality and is implemented in the algorithm, making it transparent to the user.

The particular algorithm chosen is then

$$X_1 = F_H^{-1}(U) + W$$
$$X_2 = F_{\alpha_2}^{-1}(V) \tag{7}$$

where now the subscripts can be dropped from H_1, W_1, and δ_1. Here $H = 0$ yields independent values and $H = \alpha_1$ yields random variates having the minimum possible correlation if $V = 1 - U$ and maximum correlation if $V = U$.

The set-up step reduces to satisfying the conditions

$$0 \le H \le \alpha_1$$
$$[E(F_H^{-1}(U)F_{\alpha_2}^{-1}(V)) - H\alpha_2]/(\alpha_1\alpha_2)^{1/2} = \rho \tag{8}$$

Determining the value of H satisfying conditions (8) is accomplished by bounding H and then using the modified regula falsi root finding algorithm. (See, for example, Conte and deBoor [7].) The bounds used were good enough that never more than four iterations are needed to obtain ρ within .0001. On the average two or three iterations are required. The number of integral evaluations is two more than the number of iterations.

The bounds are given in the following propositions. C_{min} and C_{max} are the minimum and maximum correlation theoretically possible, calculated from equation (3) with $H_1 = \alpha_1$, $H_2 = \alpha_2$, and $\gamma = 0$.

Proposition 4. If $\rho = C_{min}$, then $H = \alpha_1$.

Proposition 5. If $C_{min} < \rho < 0$, then $\alpha_1(\rho/C_{min})^2 < H < \alpha_1$.

Proposition 6. If $\rho = 0$, then $H = 0$.

Proposition 7. If $0 < \rho < (\alpha_1/\alpha_2)^{1/2}$, then $\alpha_1\rho^2 < H < \alpha_2$.

Proposition 8. If $\rho = (\alpha_2/\alpha_1)^{1/2}$, then $H = \alpha_2$.

Proposition 9. If $(\alpha_2/\alpha_1)^{1/2} < \rho < C_{max}$, then $\alpha_2 < H < \alpha_1(\rho/C_{max})^2$.

Proposition 10. If $\rho = C_{max}$, then $H = \alpha_1$.

The proofs follow directly from Figure 5.

4. A RELATED METHOD

In this section a method is given which will generate negatively correlated gamma variates with closed form determination of the appropriate parameter values. It will not, however, provide all possible negative correlations.

Similar to equations (1), consider

$$X_1 = Y_1 + Z + W_1$$
$$X_2 = Y_2 + Z + W_2 \qquad (9)$$

where Z, W_1, and W_2 have independent gamma distributions with shape parameters γ, δ_1 and δ_2; and Y_1 and Y_2 are Erlang k distributions obtained as

$$Y_1 = -\ln \left(\prod_{i=1}^{k} U_i\right)$$

and

$$Y_2 = -\ln \left(\prod_{i=1}^{k} (1-U_i)\right)$$

where the U_i's are independent U(0,1) random variables. The correlation of Y_1 and Y_2 is -.64493 for all values of k. Thus $\mathrm{Corr}(X_1, X_2) = (-.64493\, k + \gamma) / (\alpha_1\alpha_2)^{1/2}$.

Given α_1, α_2, and the correlation of X_1 and X_2, ρ, the appropriate parameters can be determined as follows:

$$k \leftarrow \mathrm{INT}\, (\rho(\alpha_1\alpha_2)^{1/2} / (-.64493))$$

$$\gamma \leftarrow -.64493\, k + \rho(\alpha_1\alpha_2)^{1/2}$$

$$\delta_1 \leftarrow \alpha_1 - k - \gamma$$

$$\delta_2 \leftarrow \alpha_2 - k - \gamma$$

If $k \geq 1$; and γ, δ_1 and δ_2 are nonnegative; the desired correlation is

achieved via equations (9). Otherwise the methods of the previous two sections can be used.

5. EXTENSION TO TIME SERIES MODELS

Consider the generation of a random time series X_1, X_2, ... where each X_i has a gamma(α, β) distribution and $\mathrm{Corr}(X_i, X_{i+1}) = \rho$. An autoregressive time series having these properties can be generated using the algorithm of Section 3 as follows:

ALGORITHM GTS

1. Generate $Y_1 \sim \mathrm{gamma}(\alpha, 1)$. $X_1 \leftarrow \beta Y_1$.
2. For i = 2, 3, ...,

   Generate $W_i \sim \mathrm{gamma}(\alpha - H, 1)$

   $V \leftarrow F_\alpha(Y_{i-1})$

   If $\rho < 0$, $V \leftarrow 1 - V$.

   $Y_i \leftarrow F_H^{-1}(V) + W_i$

   $X_i \leftarrow \beta Y_i$

where H is calculated to satisfy equation (8) with $\alpha_1 = \alpha_2 = \alpha$. The higher order correlations $\mathrm{Corr}(X_i, X_{i+k})$ will tend to zero as k becomes large, but the calculation of the specific values is not closed form. Extension of the algorithms of Section 2 follows in a similar manner.

APPENDIX

This appendix gives logic to determine parameter values and to generate random bivariate vectors (X_1, X_2) using the algorithm developed in Section 3. Parameters which are provided are α_1, α_2, β_1, β_2, ρ and EPS. EPS is the error allowed in ρ when determining H. Corr(H, α_1, α_2, ρ) is the correlation of X_1 and X_2 given the specified parameter values and is obtained by equation (3) which involves numerical integration of (4), (5), or (6) in Section 2 with $H_1 = H$, $H_2 = \alpha_2$, and $\gamma = 0$. Care should be taken to use an integration routine which will evaluate (4), (5), or (6) with error less than $EPS(\alpha_1\alpha_2)^{1/2}$.

ALGORITHM

Initialization (executed each time α_1, α_2 or ρ changes value)

1. (If ρ is close enough to zero, go generate independent variates.)

   If $|\rho| < EPS$, $H \leftarrow 0$ and go to step 10.

2. (Ensure that $\alpha_1 \geq \alpha_2$.)

   $IND \leftarrow 0$. If $\alpha_1 < \alpha_2$, $IND \leftarrow 1$, $T \leftarrow \alpha_1$, $\alpha_1 \leftarrow \alpha_2$, $\alpha_2 \leftarrow T$.

3. $C \leftarrow \mathrm{Corr}(\alpha_1, \alpha_1, \alpha_2, \rho)$

   If $|C| < |\rho|$, then ρ is not theoretically consistent with α_1 and α_2. Return with an error code.

4. (Bound H between A and B with associated correlation errors F and G in preparation for the modified regula falsi search in step 6.)

   If $\rho < 0$, $A \leftarrow \alpha_1(\rho/C)^2$, $B \leftarrow \alpha_1$, $G \leftarrow \rho - C$, $F \leftarrow \rho - \mathrm{Corr}(A, \alpha_1, \alpha_2, \rho)$

   If $0 < \rho < (\alpha_2/\alpha_1)^{1/2}$, $A \leftarrow \alpha_1\rho^2$, $B \leftarrow \alpha_2$, $G \leftarrow \rho - (\alpha_2/\alpha_1)^{1/2}$.

   $F \leftarrow \rho - \mathrm{Corr}(A, \alpha_1, \alpha_2, \rho)$

   If $(\alpha_2/\alpha_1)^{1/2} < \rho$, $A \leftarrow \alpha_2$, $B \leftarrow \alpha_1(\rho/C)^2$, $G \leftarrow \rho - \mathrm{Corr}(B, \alpha_1, \alpha_2, \rho)$

   $F \leftarrow \rho - (\alpha_2/\alpha_1)^{1/2}$

5. $FO \leftarrow F$

6. (Use the modified regula falsi method to determine H in steps 6-9.)

$H \leftarrow (G*A - F*B) / (G - F)$

$FN \leftarrow \rho - Corr(H, \alpha_1, \alpha_2, \rho)$

$PROD \leftarrow FO*FN$

$FO \leftarrow FN$

If $F*FN < 0$, go to step 8. Otherwise go to step 7.

7. $A \leftarrow H$, $F \leftarrow FN$. If $PROD > 0$, $G \leftarrow G/2$. Go to step 9.

8. $B \leftarrow H$, $G \leftarrow FN$. If $PROD > 0$, $F \leftarrow F/2$.

9. If $ABS(FN) > EPS$, go to step 6. Otherwise $H \leftarrow (G*A - F*B) / (G - F)$.

Generation (executed each time a new vector is needed)

10. If $H > 0$ and $H < \alpha_2$, go to step 11.

Otherwise generate $X_2 \sim \text{gamma}(\alpha_2)$, $U \leftarrow F_{\alpha_2}(X_2)$.

If $\rho < 0$, $U \leftarrow 1 - U$. $X_1 \leftarrow 0$. If $H > 0$, $X_1 \leftarrow F_H^{-1}(U)$. Go to step 12.

11. Generate $X_1 \sim \text{gamma}(H)$. $U \leftarrow F_H(X_1)$. If $\rho < 0$, $U \leftarrow 1 - U$.

$X_2 \leftarrow F_{\alpha_2}^{-1}(U)$.

12. If $H < \alpha_1$, generate $W \sim \text{gamma}(\alpha_1 - H)$ and $X_1 \leftarrow X_1 + W$.

$X_1 \leftarrow \beta_1 * X_1$, $X_2 \leftarrow \beta_2 * X_2$. If $IND = 1$, $T \leftarrow X_1$, $X_1 \leftarrow X_2$, and $X_2 \leftarrow T$.

Deliver (X_1, X_2).

A FORTRAN implementation is available upon request from the authors.

ACKNOWLEDGMENT

This research was supported by Office of Naval Research Contract N00014-77-C-0425, Code 431.

REFERENCES

1. J. H. Ahrens and U. Dieter, "Computer Methods for Sampling from Gamma, Beta, Poisson and Beta Distributions," Computing, 12, 223-246 (1974).

2. B. C. Arnold, "A Note on Multivariate Distributions with Specified Marginals," Journal of the American Statistical Association, 62, 1460-1461 (1967).

3. A. C. Atkinson, "An Easily Programmed Algorithm for Generating Gamma Random Variables," Journal of the Royal Statistical Society, A, 140, 232-234 (1977).

4. A. C. Atkinson and M. C. Pearce, "The Computer Generation of Beta, Gamma and Normal Random Variables," Journal of the Royal Statistical Society, A, 139, 431-461 (1976).

5. D. J. Best and D. E. Roberts, "Algorithm 91: The Percentage Points of the χ^2 Distribution," Applied Statistics, 24, 385-388 (1975).

6. R. C. H. Cheng, "The Generation of Gamma Variables with Noninteger Shape Parameter," Applied Statistics, 26, 71-75 (1977).

7. S. D. Conte and C. deBoor, Elementary Numerical Analysis, McGraw-Hill, New York, 1972.

8. P. J. Davis and P. Rabinowitz, Numerical Integration, Blaisdell Publishing Company, Waltham, Mass., 1967.

9. I. Deak, "The Ellipsoid Method for Generating Normally Distributed Vectors," Numerische Mathematik, forthcoming.

10. F. C. Durling, D. B. Owen, and J. W. Drane, "A New Bivariate Burr Distribution," Annals of Mathematical Statistics, 41, 1135 (1970).

11. G. S. Fishman, Concepts and Methods in Discrete Event Digital Simulation, John Wiley & Sons, New York, 1973.

12. G. S. Fishman, "Sampling from the Gamma Distribution on a Computer," Communications of the ACM, 19, 407-409 (1976).

13. M. Gargano and A. Tenenbein, "A Family of Bivariate Uniform Distributions with Application to Simulation," Graduate School of Business Administration, New York University, 100 Trinity Place, New York, NY 10006 (1978).

14. D. P. Gaver, S. S. Lavenberg, and T. G. Price, Jr., "Exploratory Analysis of Access Path Length Data for a Data Base Management System," IBM Journal for Research and Development, 449-464 (1976).

15. A. J. Greenwood, "A Fast Generator for Gamma-distributed Random Variables," COMPSTAT: Proceedings in Computational Statistics (G. Bruckman, F. Ferschl, L. Schmetterer, eds.), Physica-Verlag, Vienna, 1974.

16. R. F. Gunst and J. T. Webster, "Density Functions of the Bivariate Chi-square Distribution," Journal of Statistical Computation and Simulation, 2, 275-288 (1973).

17. J. C. Hull, "Dealing with Dependence in Risk Simulations," Operations Research Quarterly, 28, 201-213 (1977).

18. M. D. Jöhnk, "Erzeugung von Betaverteilten und Gammaverteilten Zufallszahlen," Metrika, 8, 5-15 (1964).

19. M. E. Johnson and J. S. Ramberg, "A Bivariate Distribution System with Specified Marginals," Technical Report LA-6858-MS, Los Alamos Scientific Laboratory, Los Alamos, NM 87545 (1977).

20. M. E. Johnson and A. Tenenbein, "Bivariate Distributions with Given Marginals and Fixed Measures of Dependence," Informal Report LA-7700-MS, Los Alamos Scientific Laboratory, Los Alamos, NM 87545 (1979).

21. N. L. Johnson, "Bivariates Distributions Based on Simple Translation Systems," Biometrika, 36, 297-304 (1949).

22. N. L. Johnson and S. Kotz, Distributions in Statistics: Continuous Multivariate Distributions, Wiley-Interscience, New York, 1972.

23. W. F. Kibble, "A Two-variate Gamma Type Distribution," Sankhyā, 5, 137-150 (1941).

24. A. J. Kinderman and J. F. Monahan, "Recent Developments in the Computer Generation of Student's t and Gamma Random Variables," Technical Report AMD-799, BNL-24671, Brookhaven National Laboratory, Upton, NY 11973 (Preliminary version) (1973).

25. J. F. Kottas and H. Lau, "On Handling Dependent Random Variables in Risk Analysis," Journal of the Operational Research Society, 29, 1209-1217 (1978).

26. L. Lee, "Multivariate Distributions having Weibull Properties," Technical Report, Department of Statistics, VPI & SU, Blacksburg, VA 24061 (1977).

27. G. Marsaglia, "The Squeeze Method for Generating Gamma Variates," Computers and Mathematics with Applications, 3, 321-325 (1977).

28. E. J. McGrath and D. C. Irving, Techniques for Efficient Monte Carlo Simulation. Vol. II. Random Number Generation for Selected Probability Distributions, National Technical Information Service, Springfield, VA, 1973.

29. K. V. Mardia, Families of Bivariate Distributions, Griffon, London, 1970.

30. G. A. Mihram and A. R. Hultquist, "A Bivariate Warning-time/Failure-time Distribution," Journal of the American Statistical Association, 62, 589-599 (1967).

31. P. A. P. Moran, "Testing for Correlation between Non-negative Variates," Biometrika, 54, 385-394 (1967).

32. R. J. Polge, E. M. Holliday, and B. K. Bhagavan, "Generation of a Pseudo-random Set with Desired Correlation and Probability Distribution," Simulation, 20, 153-158 (1973).

33. G. Ronning, "A Simple Scheme for Generating Multivariate Gamma Distributions with Non-negative Covariance," Technometrics, 19, 179-183 (1977).

34. E. M. Scheuer and D. S. Stoller, "On the Generation of Normal Random Vectors," Technometrics, 4, 278-281 (1962).

35. B. W. Schmeiser and A. I. Ali, "The n-dimensional Polar Method for Generating Pseudo-random Normal Deviates," IEOR Technical Report 77011, Southern Methodist University, Dallas, TX 75275 (1977).

36. B. W. Schmeiser and R. Lal, "Squeeze Methods for Generating Gamma Variates," OREM Technical Report 78009, Southern Methodist University, Dallas, TX 75275 (1978).

37. P. R. Tadikamalla, "Computer Generation of Gamma Random Variables," Communications of the ACM, 21, 419-422 (1978a).

38. P. R. Tadikamalla, "Computer Generation of Gamma Random Variables-II," Communications of the ACM, 21, 925-927 (1978b).

39. K. Takahasi, "Note on the Multivariate Burr's Distribution," Annals of the Institute of Statistical Mathematics, 17, 257-260 (1965).

40. C. S. Wallace, "Transformed Rejection Generators for Gamma and Normal Pseudo-random Variables," The Australian Computer Journal, 8, 103-109 (1976).

41. N. D. Wallace, "Computer Generation of Gamma Random Variables with Non-integral Shape Parameters," Communications of the ACM, 17, 691-695 (1974).

42. J. Whittaker, "Generating Gamma and Beta Random Variables with Non-integral Shape Parameters," Applied Statistics, 23, 210-214 (1974).

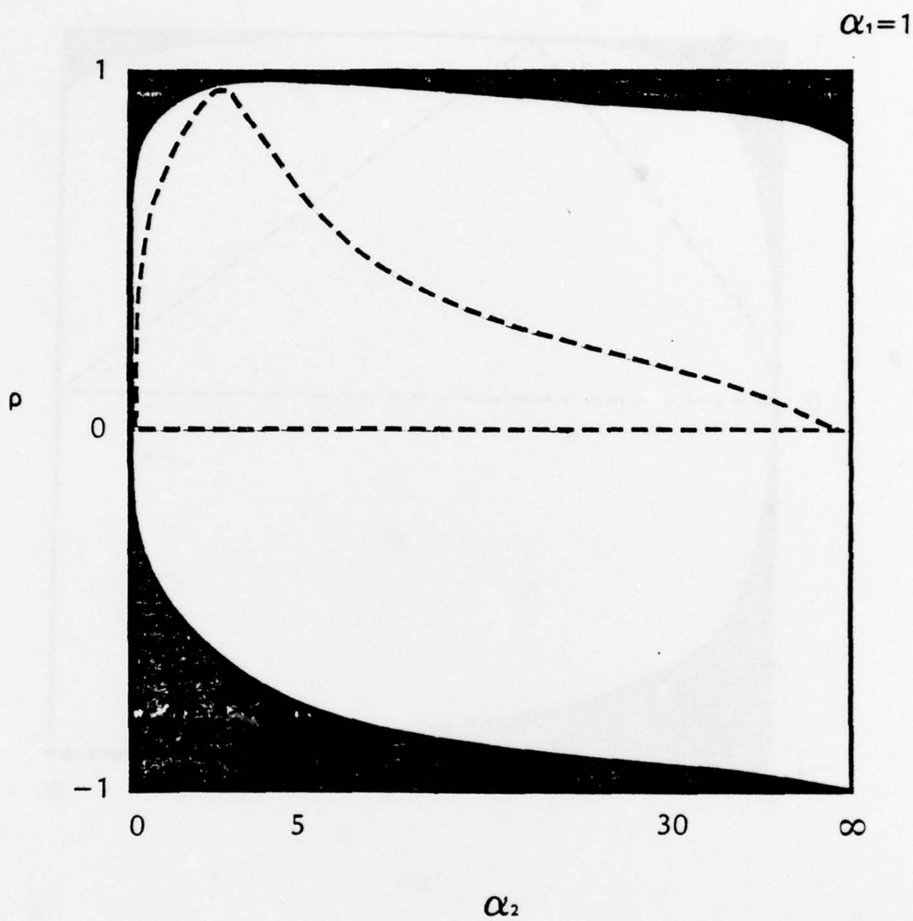


Figure 1(a). Theoretical limits for bivariate gamma correlations for $\alpha_1 = 1$.

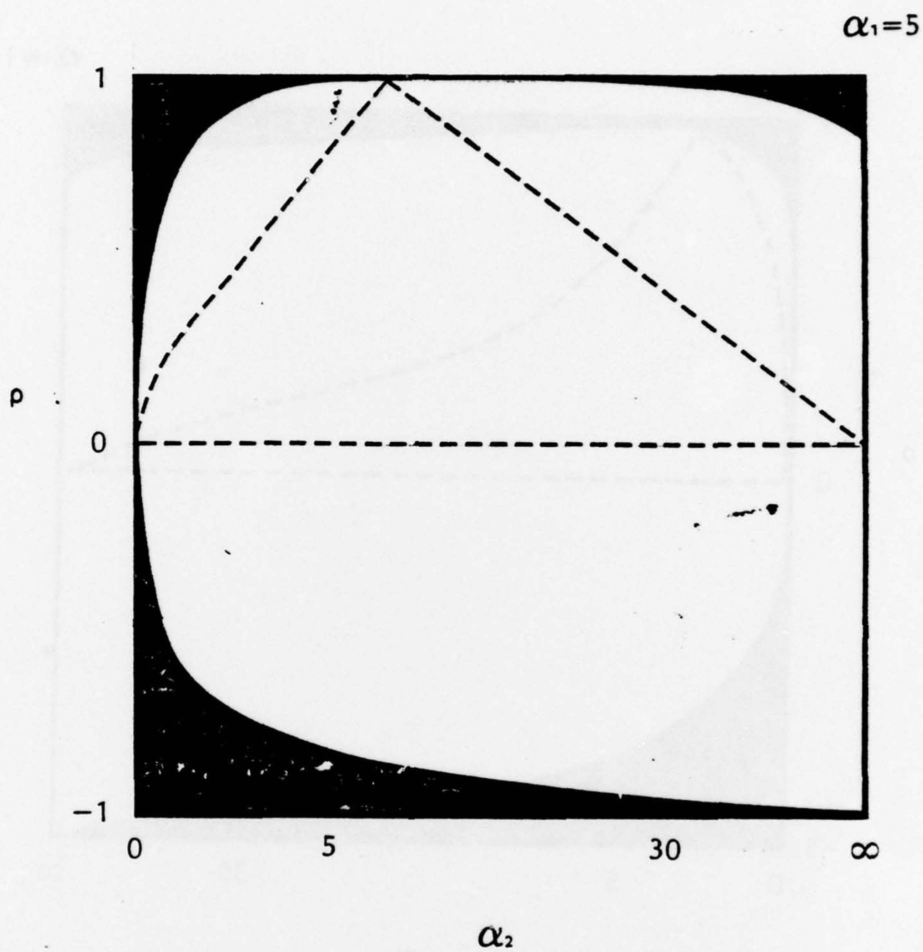


Figure 1(b). Theoretical limits for bivariate gamma correlations for $\alpha_1 = 5$.

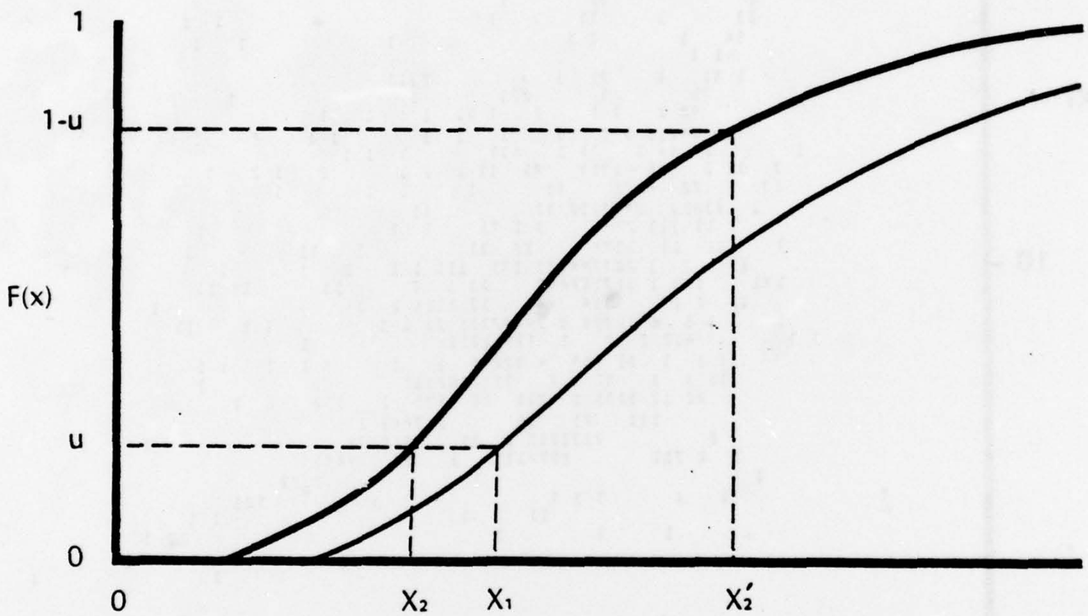


Figure 2. The relationship between two random variables having minimum and maximum correlation.

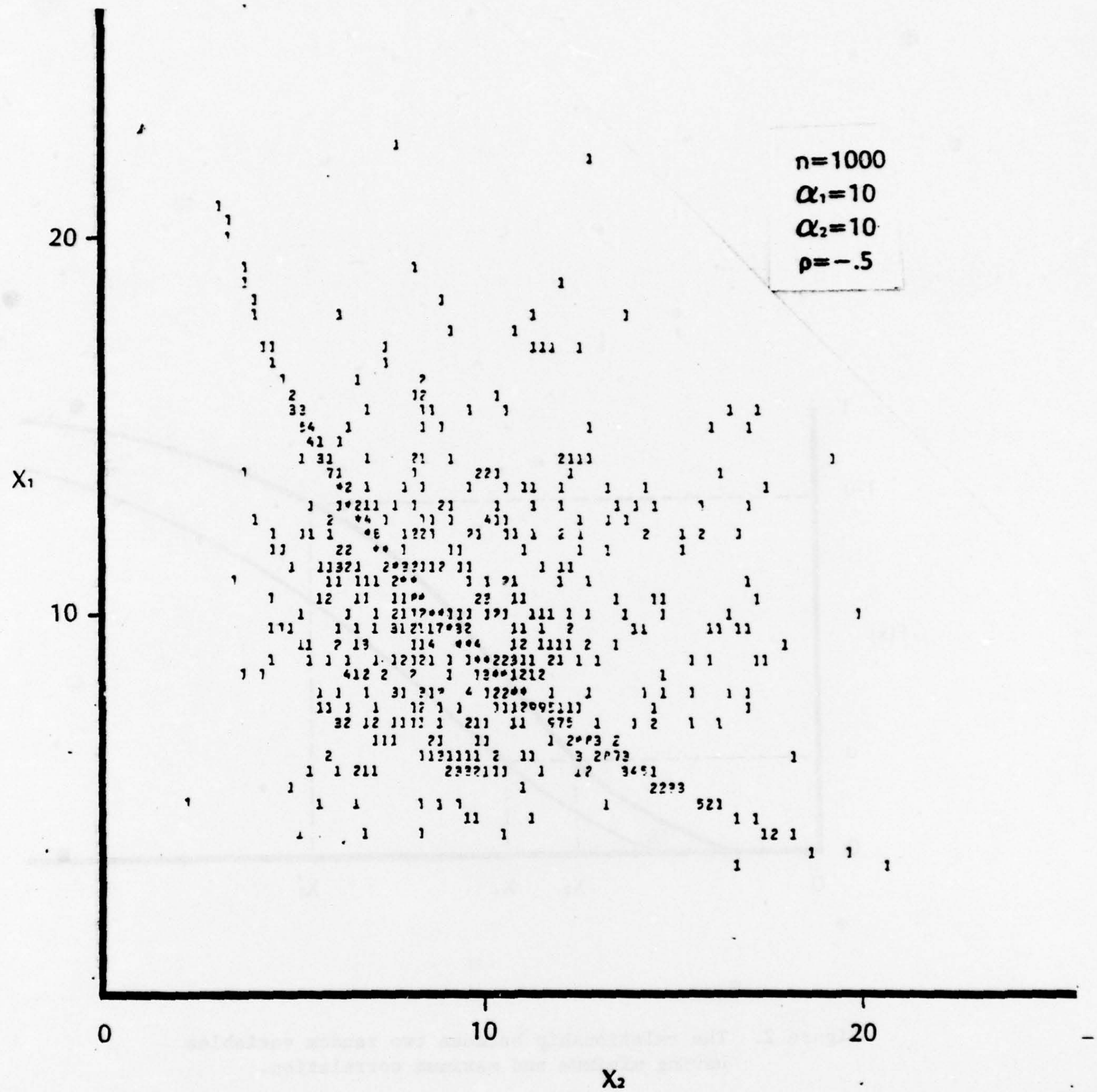


Figure 3. A scatter plot of bivariate gamma points using the composition algorithm GCOMP.

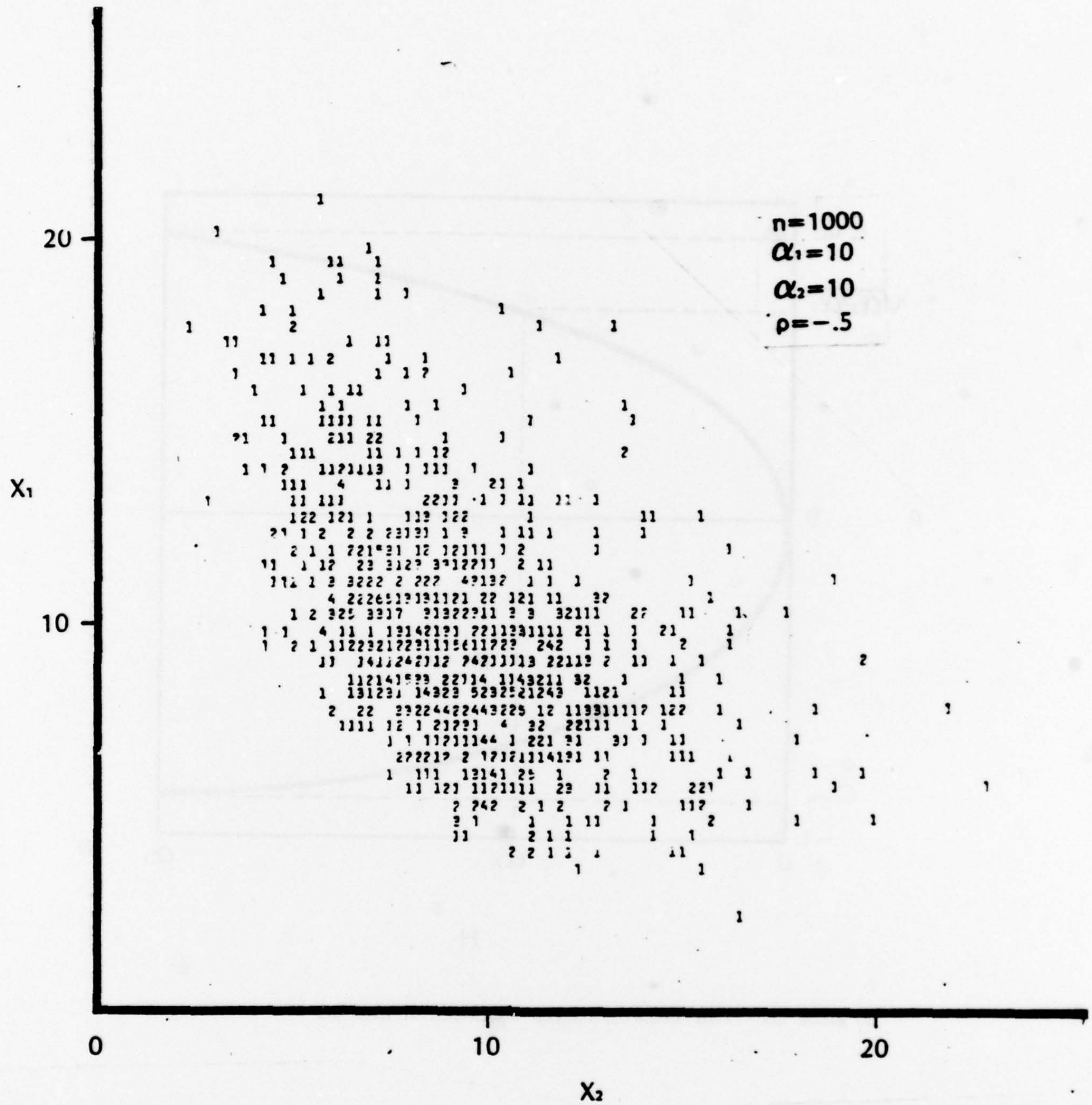


Figure 4. A scatter plot of bivariate gamma points using the algorithm of Section 3, given in the Appendix.

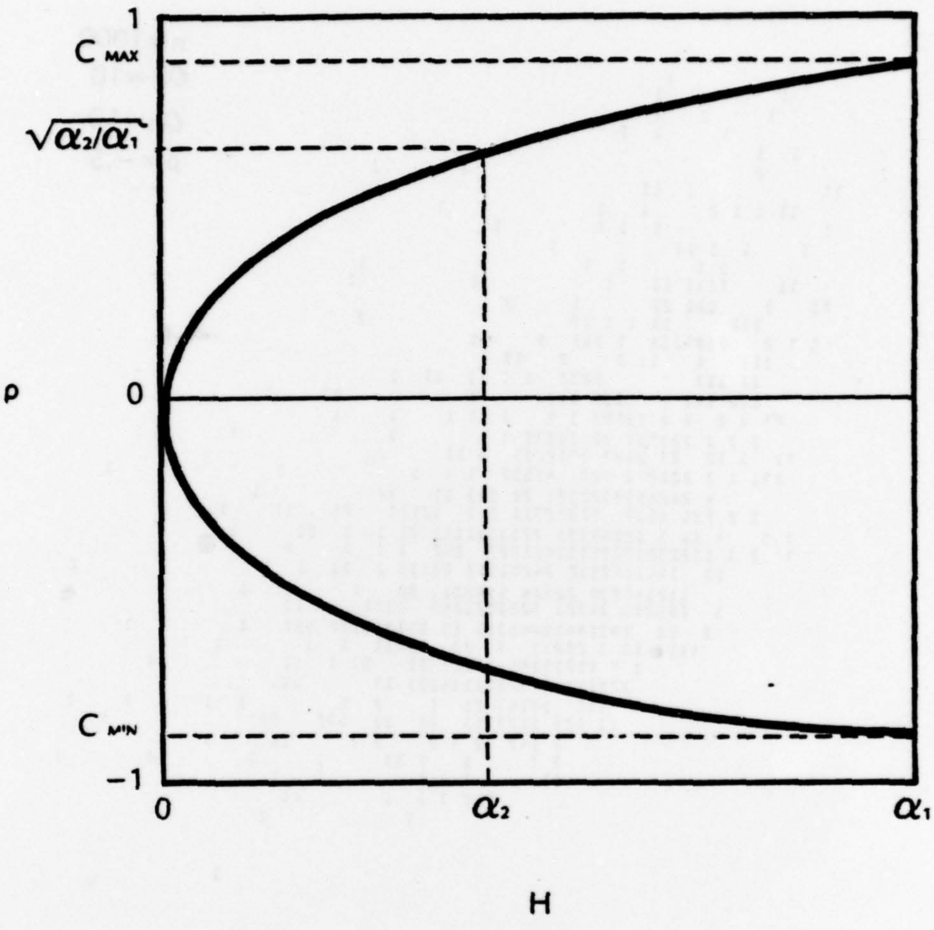


Figure 5. The relationship between H and ρ.

APPENDIX II

Fortran implementation of the algorithm developed in Section 3, including a driver program and numerical integration subroutine.

```
      PROGRAM MAIN(INPUT,OUTPUT,TAPE5=INPUT,TAPE6=OUTPUT)
C     BRUCE SCHMEISER AND RAM LAL     6/79    SOUTHERN METHODIST UNIVERSITY
C     MAIN PROGRAM TO TEST BIVARIATE GAMMA GENERATORS
      DIMENSION AL(4),R(9)
      DATA AL/.1,1.,5.,100./
      DATA R/-.8,-.5,-.1,-.1E-5,0.,.1E-5,.1,.5,.8/
      N = 10
      DO 10 I=1,4
      DO 10 J=1,I
      DO 10 K=1,9
      A1 = AL(I)
      A2 = AL(J)
      B1=1.
      B2=1.
      RHO=R(K)
      WRITE (6,2) A1,A2,B1,B2,RHO
    2 FORMAT (//10(≠ ..........≠)/≠ ..........≠,5F15.6,≠.............≠)
      SX=SY=SX2=SY2=SXY=0.
      TIME = SECOND(X)
      DO 100 L = 1,N
      CALL BGAM(A1,B1,A2,B2,RHO,.0001,X,Y,IER)
      IF (IER .NE. 0)   GO TO 10
      SX = SX + X
      SX2 = SX2 + X*X
      SY = SY + Y
      SY2 = SY2 + Y*Y
  100 SXY = SXY + X*Y
      TIME = (SECOND(X) - TIME) * 1000. / N
      AX = SX / N
      AY = SY / N
      VX = SX2/N - AX*AX
      VY = SY2/N - AY*AY
      CORR = (SXY/N - AX*AY) / SQRT(VX*VY)
      WRITE (6,1) A1,A2,B1,B2,RHO,AX,AY,VX,VY,CORR,N,TIME,IER
    1 FORMAT (10F8.3,I5,F8.3,I5)
   10 CONTINUE
  999 RETURN
      END
```

```
      SUBROUTINE BGAM(A1,B1,A2,B2,RHO,EPS,X,Y,IER)
C     BRUCE SCHMEISER AND RAM LAL  6/79  SOUTHERN METHODIST UNIVERSITY
C     TO GENERATE ONE BIVARIATE VECTOR (X,Y) HAVING MEAN VALUES
C        A1*B1 AND A2*B2 AND CORRELATION WITHIN EPS OF RHO.
C     THE METHOD IS  X = Z1 + W AND Y = Z2, WHERE CORR(Z1,Z2) IS
C        MAXIMIZED IF RHO .GT. 0 AND IS MINIMIZED IF RHO .LT. 0.
C     REFERENCE BRUCE SCHMEISER AND RAM LAL "GENERATION OF BIVARIATE
C        GAMMA RANDOM VECTORS", OREM 79009, DEPARTMENT OF OPERATIONS
C        RESEARCH AND ENGINEERING MANAGEMENT, SOUTHERN METHODIST
C        UNIVERSITY, DALLAS, TX  75275.
      DATA SA1/-1./
      IF (A1 .EQ. SA1 .AND. A2 .EQ. SA2 .AND. RHO .EQ. SRHO .AND.
     1    EPS .EQ. SEPS)  GO TO 100
C
C*****NEW PARAMETERS, SO DETERMINE NEW CONSTANTS************************
C
      SA1 = A1
      SA2 = A2
      SRHO = RHO
      SEPS = EPS
      AL1 = A1
      AL2 = A2
      IF (A1 .GT. A2)  GO TO 10
      AL1 = A2
      AL2 = A1
C
C*****CHECK FOR VALID PARAMETERS
C
   10 IER = 1
      IF (A1 .LE. 0. .OR. A2 .LE. 0. .OR. EPS .LE. 0.)  RETURN
      IER = 2
      S = SQRT(A1*A2)
      CMAX = (EXY(A1,A2,RHO) - A1*A2) / S
      IF (ABS(RHO) .GT. ABS(CMAX)) RETURN
      IER = 0
C
C*****PARTITION AL1 INTO H AND AL1-H TO OBTAIN THE DESIRED CORRELATION.
C     FIRST BOUND THE SOLUTION.
C
      H = 0.
      IF (ABS(RHO) .LT. EPS)  GO TO 100
      IF (RHO .GT. 0.)  GO TO 20
      A = AL1 * (RHO/CMAX)**2
      B = AL1
      E = RHO - (EXY(A,AL2,RHO) - A*AL2) / S
      G = RHO - CMAX
      GO TO 30
   20 IF(RHO .GT. SQRT(AL2/AL1))  GO TO 25
      A = AL1 * RHO * RHO
      B = AL2
      E = RHO - (EXY(A,AL2,RHO) - A*AL2) / S
      G = RHO - SQRT(AL2/AL1)
      GO TO 30
   25 A = AL2
      B = AL1 * (RHO/CMAX)**2
      E = RHO - SQRT(AL2/AL1)
```

```
      G = RHO - (EXY(B,AL2,RHO) - B*AL2) / S
   30 FO = F
C
C     APPLY THE MODIFIED REGULA FALSI ROOT FINDING METHOD TO FIND
C     H, GIVEN THE INITIAL VALUES DETERMINED ABOVE.
C

   40 H = (G*A - F*B) / (G-F)
      FN = RHO - (EXY(H,AL2,RHO) - H*AL2) / S

      PROD = FO*FN
      FO = FN
      IF (F*FN .LT. 0.)  GO TO 50
      A = H
      F = FN
      IF (PROD .GT. 0.)  G = G/2.
      GO TO 60
   50 B = H
      G = FN
      IF (PROD .GT. 0.)  F = F/2.
   60 IF (ABS(FN) .GT. EPS)  GO TO 40
      H = (G*A - F*B) / (G-F)


C
C*****GENERATE THE RANDOM VECTOR (X,Y).**************************************
C
  100 IF(H .LT. AL2 .AND. H .NE. 0.) GO TO 150
      Y = RGAM(AL2,ISEED)
      CALL GMCDFN(Y,AL2,U,IER)
      IF (RHO .LT. 0.)  U = 1.-U
      X = 0.
      IF (H .GT. 0.)  X = GINV (U,H)
      GO TO 180
  150 X = RGAM(H,ISEED)
      CALL GMCDFN(X,H,U,IER)
      IF(RHO .LT. 0.) U = 1.-U
      CALL GINV(U,AL2)
  180 IF (H .LT. AL1)  X = X + RGAM(AL1-H,ISEED)
      X = B1*X
      Y = B2*Y
      IF (A1 .EQ. AL1)  GO TO 200
      T = X
      X = Y
      Y = T
  200 RETURN
      END
```

NOTE: RGAM generates a gamma variate.
GINV (U,H) = $F^{-1}_{H}(U)$.

```
      FUNCTION FXY(AL,AB,IPS)
C
C     THIS FUNCTION EVALUATES AN INTEGRAL USING GAUSSIAN TWENTY-FOUR
C     POINT FORMULA. IT CALLS FUNCTION FF(WHOSE INTEG.IS BEING EVALUATED)
C     WITH REQUISITE PARAMETERS.
C
      REAL IPS
      A=.5
      C=.49759360999851068
      Y=.006170614899993599*(FF(A+C,AL,AB,IPS)+FF(A-C,AL,AB,IPS))
      C=.48736427798565475
      Y=Y+.014265694314466832*(FF(A+C,AL,AB,IPS)+FF(A-C,AL,AB,IPS))
      C=.469137276001366380
      Y=Y+.022138719408709903*(FF(A+C,AL,AB,IPS)+FF(A-C,AL,AB,IPS))
      C=.44320776350220052
      Y=Y+.029649292457718390*(FF(A+C,AL,AB,IPS)+FF(A-C,AL,AB,IPS))
      C=.410000992986951460
      Y=Y+.036673240705540153*(FF(A+C,AL,AB,IPS)+FF(A-C,AL,AB,IPS))
      C=.370062095789277180
      Y=Y+.043095080765976683*(FF(A+C,AL,AB,IPS)+FF(A-C,AL,AB,IPS))
      C=.32404682596848778
      Y=Y+.048809326052056944*(FF(A+C,AL,AB,IPS)+FF(A-C,AL,AB,IPS))
      C=.27271073569441977
      Y=Y+.053722135057982817*(FF(A+C,AL,AB,IPS)+FF(A-C,AL,AB,IPS))
      C=.21689675381302257
      Y=Y+.057752834026862801*(FF(A+C,AL,AB,IPS)+FF(A-C,AL,AB,IPS))
      C=.15752133984808169
      Y=Y+.060835236463901696*(FF(A+C,AL,AB,IPS)+FF(A-C,AL,AB,IPS))
      C=.09555943373680815
      Y=Y+.062918728173414148*(FF(A+C,AL,AB,IPS)+FF(A-C,AL,AB,IPS))
      C=.03202844643130281 3
      FXY = Y+.063969097673376078*(FF(A+C,AL,AB,IPS)+FF(A-C,AL,AB,IPS))
      RETURN
      END

      FUNCTION FF(U,A1,A2,RHO)
      IF (RHO .GT. 0.)  FF = GINV(U,A1) * GINV(U,A2)
      IF (RHO .LE. 0.)  FF = GINV(U,A1) * GINV(1-U,A2)
      RETURN
      END
```

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER OR 79009	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) COMPUTER GENERATION OF BIVARIATE GAMMA RANDOM VECTORS		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Bruce W. Schmeiser and Ram Lal		8. CONTRACT OR GRANT NUMBER(s) ONR N00014-77-C-0425
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Operations Research Southern Methodist University Dallas, TX 75275		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 277-236
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Analysis Program, Code 431 Office of Naval Research Arlington, VA 22217		12. REPORT DATE July 1979
		13. NUMBER OF PAGES 36
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution of this document is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Monte Carlo	Gamma Distribution
Distribution Sampling	Bivariate Gamma
Simulation	Multivariate Gamma

20. The use of bivariate gamma distributions in simulation modeling is considered. A family of algorithms is given, any member of which can be used to generate random bivariate vectors having any gamma marginal distributions and any theoretically possible correlation both positive and negative. Computational considerations are discussed, including the selection of parameters to provide regression curves consistent with the modeler's needs. A modification which is easier to implement and provides most, but not all, correlations is also given. Finally the use of these algorithms to generate first order autoregressive time series is discussed.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)